ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

# Research on Low-Light Image Enhancement Algorithm Based on Dual-Branch Structure

**Xiaoting Niu*, Guanghai Zheng**

*School of Software, Dalian Jiaotong University, Dalian, China*
**Corresponding author: 905922247@qq.com*

***Abstract:*** *This study addresses the prevalent issues of color bias and detail loss in low-light images, which significantly affect image quality. A low-light enhancement algorithm based on the combination of convolutional neural networks (CNN) and a dual-branch structure is proposed to address the problems of detail loss and color shift that existing single-stage CNN models cannot effectively handle. This module is trained with a contrastive regularization method on the basis of the dual-branch structure to ensure the consistency of detail distribution between the generated images and reference images. In terms of the loss function, a color loss function $L_{color}$ is added to balance the low-light enhancement and color bias issues. Experimental results show that the pre-trained model achieved a peak signal-to-noise ratio (PSNR) of 22.133 dB and a structural similarity (SSIM) of 0.873 on the LOL dataset. This indicates that the proposed algorithm significantly improves the balance between image brightness and detail retention, and performs well in enhancing the brightness of natural low-light scene images.*

***Keywords:*** *CNN; Dual-Branch Structure; Contrastive Regularization; Low-Light Enhancement; Residual Learning*

## 1. Introduction

In recent years, with the rapid development of deep learning technology, convolutional neural networks (CNNs) have achieved significant results in the field of image processing. Image enhancement is an important subfield of image processing, often serving as a preprocessing step to prepare for subsequent image analysis and processing tasks. The aim of image enhancement is to improve image brightness, adjust image contrast, recover details hidden in the dark, and increase the utility value of images through corresponding technical means. CNNs possess strong feature extraction and adaptive learning capabilities, enabling them to automatically learn complex patterns and relationships within images, thus providing new solutions for low-light image enhancement. Therefore, how to effectively enhance the quality of low-light images using CNNs to make them closer to the visual effects under natural lighting has become an important topic in the field of image processing.

Image de-noising methods based on image filtering are widely used [1] and perform best when dealing with all-band noise, adaptive selection dependent on noise level, combined with other de-noising techniques, and specific types of noise (such as Gaussian noise) [2]. However, there are obvious limitations in dealing with complex noise and maintaining image detail. However, using learn-based methods, such as DnCNN, FFDNet, etc., denoising is realized by learning the mapping from noisy image to clean image [3-5], but there are also some challenges, especially in terms of data requirements, computing resources and model generalization ability. In the actual process of improving image quality, the enhancement of light is accompanied by the expansion of noise signal, and the removal of noise signal will often blur the features of dark light image. In recent years, with the rapid development of deep learning, the method based on convo-lutional neural network (CNN) has been widely used in image processing, and has also achieved unprecedented achievements in dark light enhancement and image denoising [6-9]. The real-time low-light enhancement algorithm (Zero-DCE) proposed by Guo C et al., which does not require reference data, uses neural networks to fit a brightness mapping curve, and then generates a brightening image based on the curve and the original image [9]. However, Zero-DCE can only achieve brightness enhancement and does not work on image deblurring. Lin L et al proposed a deep unfolding network based on Retinex (URetinex-Net), which decomposed low-light images into reflection layer and illumination layer, adaptive fitting implicit priori in a data-driven way, and realized noise suppression and detail retention of the final

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

decomposition results [10]. However, although this deep learning method can achieve the adjustment of the overall brightness, it lacks the suppression of degraded information such as noise. Later, Wriza W et al used URetinex-Net and TRBA to enhance low-light image in license plate recognition on this basis, and this method had a good effect on improving the environmental accuracy of night license plate recognition system [11]. Lin S et al. proposed a synchronous multi-scale dark light enhancement network (SMNet) method, which learns feature streams of different scales in a top-down to bottom-up manner through multiple L&G modules in series to achieve image denoising. This method has good effects on dark light image enhancement and noise suppression in natural scenes [12].

In this paper, a dual-branch network is constructed, and contrastive learning regularization is introduced on the basis of the dual-branch structure to enhance the model's processing of image details and structure. A color preservation function is subsequently added to effectively address the issues of color bias and detail loss. Furthermore, due to the diversity of the data space in contrastive learning, the process of pulling the restored results closer to normal light images and pushing them away from low-light images causes a shift in color. Therefore, a color loss function $L_{color}$ is added on this basis to balance the enhancement of low-light images and the issue of color bias.

## 2. Design of Low-Light Image Enhancement Model Based on Dual-Branch Structure

### *2.1 Model Construction*

Figure 1 shows the overall framework of the Dual-Branch Structure model[13].

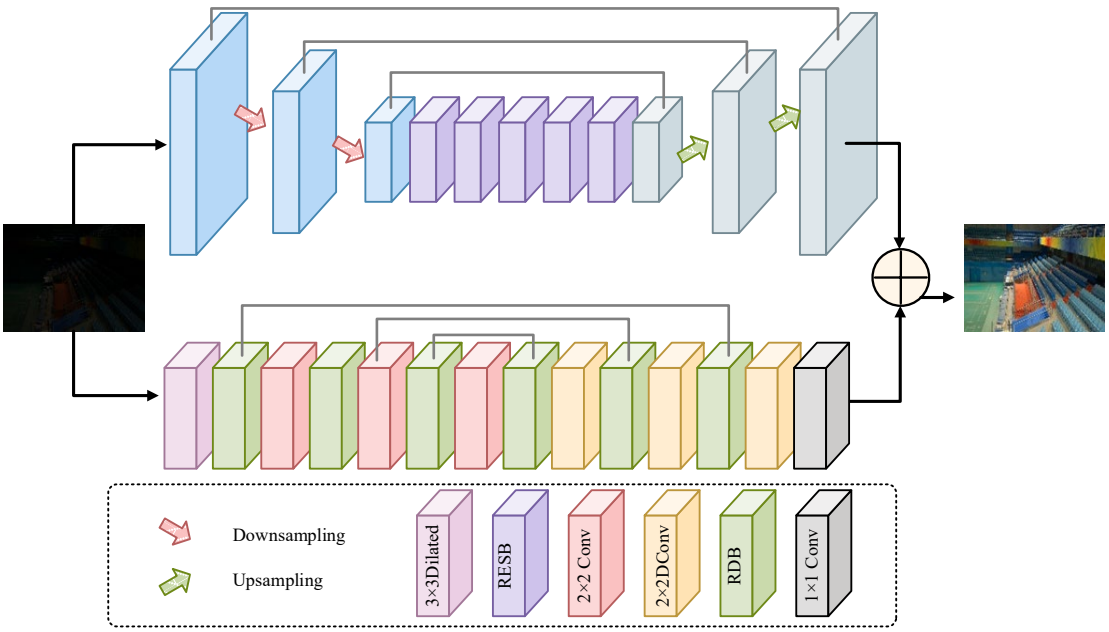



*Figure 1. Overall framework of Dual-Branch Structure Model*

The upper branch of the dual-branch model first uses downsampling methods to extract deep features, processes these deep features through multiple layers of residual blocks RESB(Figure 2), and then uses upsampling to reconstruct and restore the low-light image. The lower branch, addressing issues such as detail loss and color shift that are prone to occur in low-light enhancement, first uses smoothing convolution to smooth the image, reducing noise, blurring edges, or enhancing certain characteristics of the image, preparing for subsequent feature extraction. Secondly, it utilizes residual dense blocks RDB(Figure 3) with convolutional layers in skip connections, effectively solving the problem of gradient vanishing in deep neural networks, achieving the goal of preserving color and detail information.

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

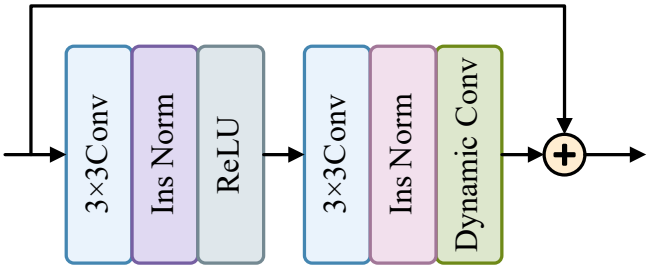



*Figure 2. RESB Residual Block*

Each RESB consists of two 3×3 convolutional layers, two instance normalization layers, a ReLU activation function, and a dynamic convolution layer. Skip connections are introduced to address the vanishing gradient problem in the training of deep networks. Dynamic convolution is also incorporated to adaptively adjust convolution parameters based on the input image, further enhancing the network's performance in low-light enhancement[14]. This new design increases the model's expressive power without increasing the network's depth or width.

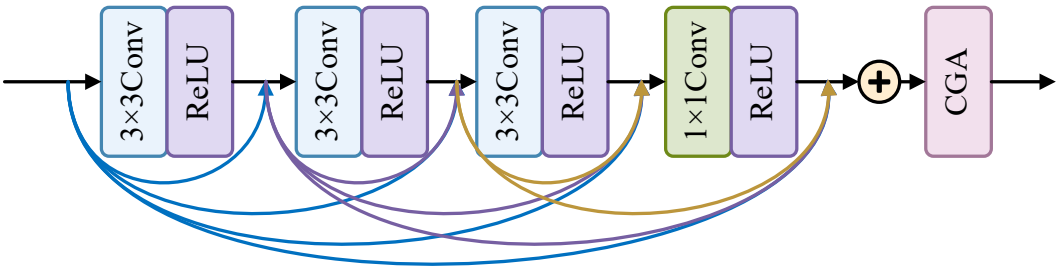



*Figure 3. RDB Residual Block*

As shown in Figure 3, the Residual Dense Block (RDB) is composed of three pairs of convolutional layers with ReLU activation functions. After dimension reduction by a 1×1 convolution kernel, the convolutional layers are connected through a dense structure, and then residually connected with the input feature map. The characteristics of the residual dense network allow the multi-scale features extracted by the earlier layers to be fully utilized. This design not only enhances the model's ability to capture details and color information but also prevents the vanishing gradient problem. Finally, a Content-Guided Attention (Figure 4) block is introduced, enabling the model to focus more on the important information in each channel of the image and ignore irrelevant information, thereby improving the effect of low-light enhancement.

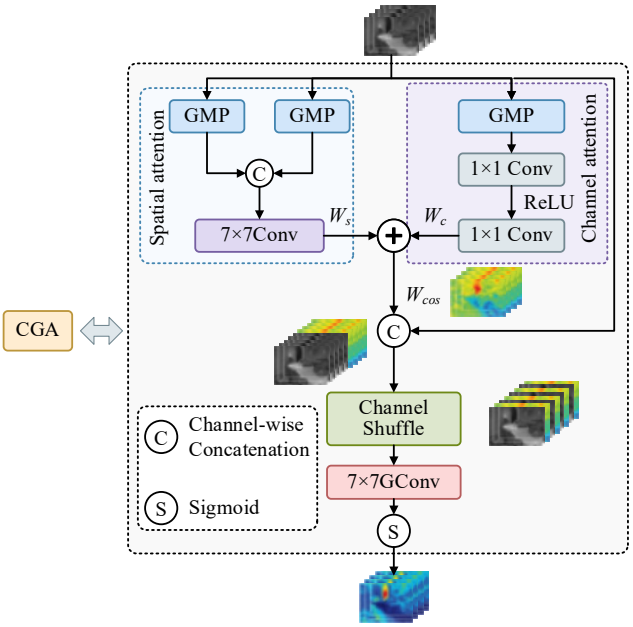



*Figure 4. CGA Workflow Diagram*

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

Figure 4 shows the Content-Guided Attention (CGA)[15]. A coarse-to-fine processing procedure aimed at generating an attention map with channel-wise importance. Initially, a coarse spatial attention map is produced, which is then refined according to each channel of the input feature map to yield the final spatial attention map. By using the content of the input features to guide the generation of the attention map, CGA focuses more on the unique feature parts of each channel, which allows for better recalibration of features and learning channel-specific attention maps that focus on the differences between channels.

$X$ represents the input features, first calculate $W_c$ and $W_s$ separately.

$$W_c = Conv_{1\times1}(\mathrm{Re}\,LU(conv_{1x1}(X^c_{GAP}))) \tag{1}$$

$$W_s = Conv_{7\times7}(contact(X^s_{GAP}, X^s_{GMP})) \tag{2}$$

$\mathrm{C}_{k\times k}(\bullet)$ represents a convolutional layer with a kernel size of $K\times K$ ,contact indicates channel-wise concatenation, $X^c_{GAP}$ , $X^s_{GAP}$ , $X^s_{GMP}$ represent the features obtained from global average pooling operations across spatial dimensions, global average pooling operations across channel dimensions, and global max pooling operations across channel dimensions, respectively. The two 1x1 convolutions are used to reduce channel dimensions and restore channel dimensions to limit model complexity.

Then combine $W_c$ and $W_s$ to obtain the initial content-guided attention:

$$\hat{W} = W_s + W_c \tag{3}$$

Next, $\hat{W}$ is concatenated with the input features through channel shuffling, allowing it to adjust the channels based on the input features. The specific operation is as follows:

$$W = sigmoid(Gc_{7\times7}(shuffle(contact(X,\hat{W})))) \tag{4}$$

Among them, $Gc_{7\times7}(\bullet)$ represents a 7×7 grouped convolution.

The CGA block assigns a unique Spatial Importance Map (SIM) to each channel, guiding the model to focus on more useful information encoded in the features. Generating channel-specific SIMs in a content-guided manner enhances the model's ability to extract features.

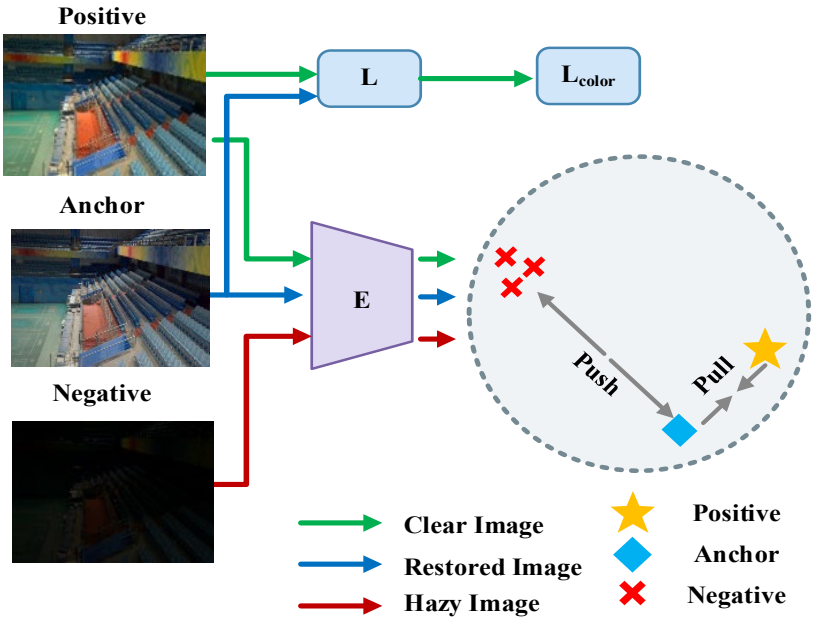



*Figure 5. Structure of Contrastive Learning Diagram*

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

Figure 5 illustrates the flowchart of the contrastive regularization technique based on the idea of contrastive learning[16]. As can be seen from the figure, contrastive regularization optimizes the model's feature representation by bringing similar samples closer and pushing dissimilar samples further apart.

Furthermore, due to the diversity of the data space in contrastive learning, the process of pulling the restoration results closer to normal light images and pushing them away from low-light images causes a shift in color. Therefore, a color loss function $L_{color}$ is added on this basis.

### *2.2 Parameter Setting*

The environment and configuration parameters used for the training and testing of the experimental network model are shown in Table 1.

*Table 1 Experimental Environment Configuration*

| | |
|---|---|
| Experiment Platform | Windows 10 Operating System |
| Processor | Intel(R)Core(TM)i9-10900X |
| GPU | NVIDIA GeForce RTX 3090 |
| Memory | 64GB |
| Programming Language | Python 3.9 |
| Programming Framework | PyTorch |

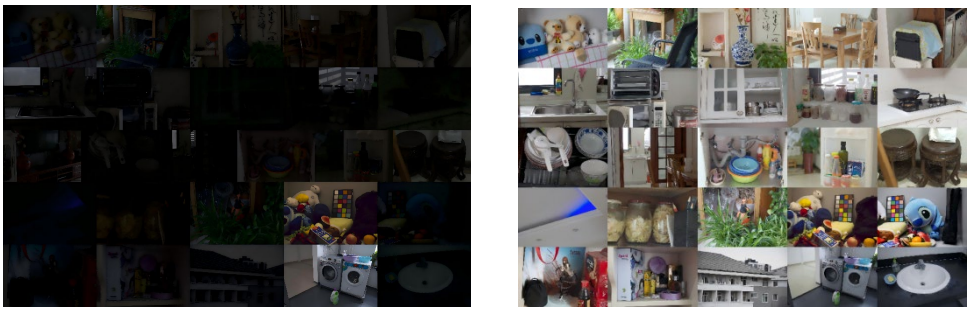
In the experiment, the learning rate is 0.0001, and the decay rate is 0.1. The Adam optimizer is used for optimization, with a batch size of 4.

## 3. Model Training

### *3.1 Data Set Preparation*

Considering that low-light image enhancement requires a large number of paired low-light and clear images as training data, but it is difficult to obtain real reference images to form a training set, the training dataset used in this experiment includes the LOL dataset, which is a mainstream real-world low-light image dataset, the GladNet dataset from the GladNet method, and the LSRW dataset.

The LOL dataset consists of 500 pairs of low-light and normal-light images, including 485 training pairs and 15 testing pairs. As shown in Figure 6, the LOL dataset includes images under various lighting conditions in different scenes.



(a) (b)

*Figure 6. LoL dataset examples (a) Low-light image (b) Normal-light image*

### *3.2 Loss Function*

In order to better restore the original colors of low-light images, this paper uses the $L_C$ color loss function as the loss function for this experiment.

$$L = \min \left\| J - \phi(I, \omega) \right\| + \beta \cdot \rho(G(I), G(J), G(\phi(I, \omega))) \tag{5}$$

Derivation:

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

$$L = \min \left\| J - \phi(I,\omega) \right\|_1 + \beta \sum_{i=1}^{n} \omega_1 \cdot \frac{D(G_i(J), G_i(\phi(I,\omega)))}{D(G_i(I), G_i(\phi(I,\omega)))} \tag{6}$$

In which, Gi, i=1,2...nrepresents the i-th hidden feature extracted from a fixed pre-trained model. D(x,y) is the L1 distance between x and y. $\omega_1$ is a weighting coefficient. Due to the nature of contrastive learning, it tends to cause color shifts when pulling the restored results closer to normal-light images and pushing them away from low-light images, as a result of the diversity in the data space. Therefore, a color preservation function is added on this basis, as shown in Equation (7):

$$L_{color} = \sum_{\forall (p,q) \in \varepsilon} (J^P - J^q)^2, \varepsilon = \{(R,G),(R,B),(G,B)\} \tag{7}$$

Finally, the total color loss function, as shown in Equation (8):

$$L_C = L + L_{color} \tag{8}$$

## 4. Experimental Results and Analysis

The proposed algorithm was tested in three datasets respectively, and the test results were compared with other algorithms. The comparison algorithms include: ZERO (Z-Score Normalization)[9], URetinex(a deep expansion network based on Retinex theory), ILL (Illumination-Adaptive Transformer network), SMNet (Synchronous Multi-Scale Dark Light Enhancement Network)[12].

Figure 7 shows the recovery results of different algorithms on the LOL dataset.

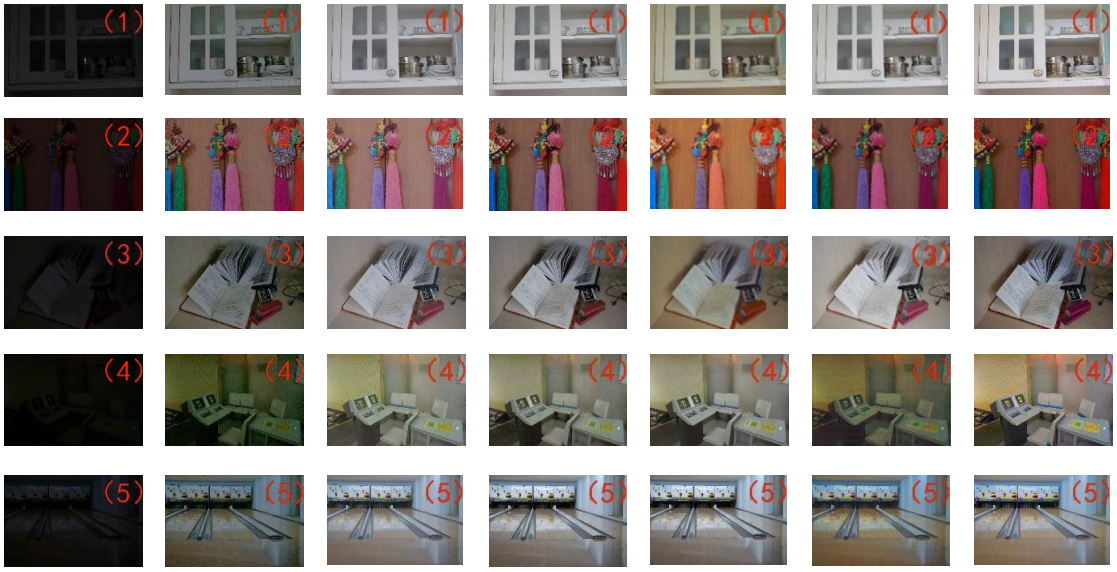



*Figure 7. Restoration results of the same algorithm on the LOL dataset*

Through comparison, it was found that the ZERO algorithm does not significantly enhance the brightness of the restored images, has an overall cool color tone, and loses some image details, as seen in the cabinet contents of Figure (b)(1). This may be due to the curves in the ZERO method not being monotonically increasing, leading to areas that were originally brighter becoming darker. From Figure (c), it can be observed that the URetinex method improves image brightness and contrast, but it also introduces some artifacts or unnatural visual effects. Figures (c)(2) and (c)(3) show color distortion due to over-enhancement, and the handling of details appears unnatural. It can be seen from the figures that the ILL method fails to adequately enhance image contrast, and the processed images have a generally low color temperature and contain artifacts, such as the ground area in Figure (d)(5), with overall low saturation. The SMNet method can effectively increase image brightness, but the color contrast of the processed images is generally low, which can easily lead to color distortion. Figure (e)(1) has a yellowish tone, further illustrating the color distortion issue caused by SMNet. Upon observation and comparison, the proposed algorithm demonstrates better image enhancement effects in the figures,

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

effectively increasing image brightness and contrast while maintaining clarity and naturalness of image details. Compared to the performance of other methods on the LOL dataset, the proposed algorithm may be more effective in processing low-light images and more in line with human visual perception.

Figure 8 shows the restoration results of different algorithms on the GladNet-Dataset .

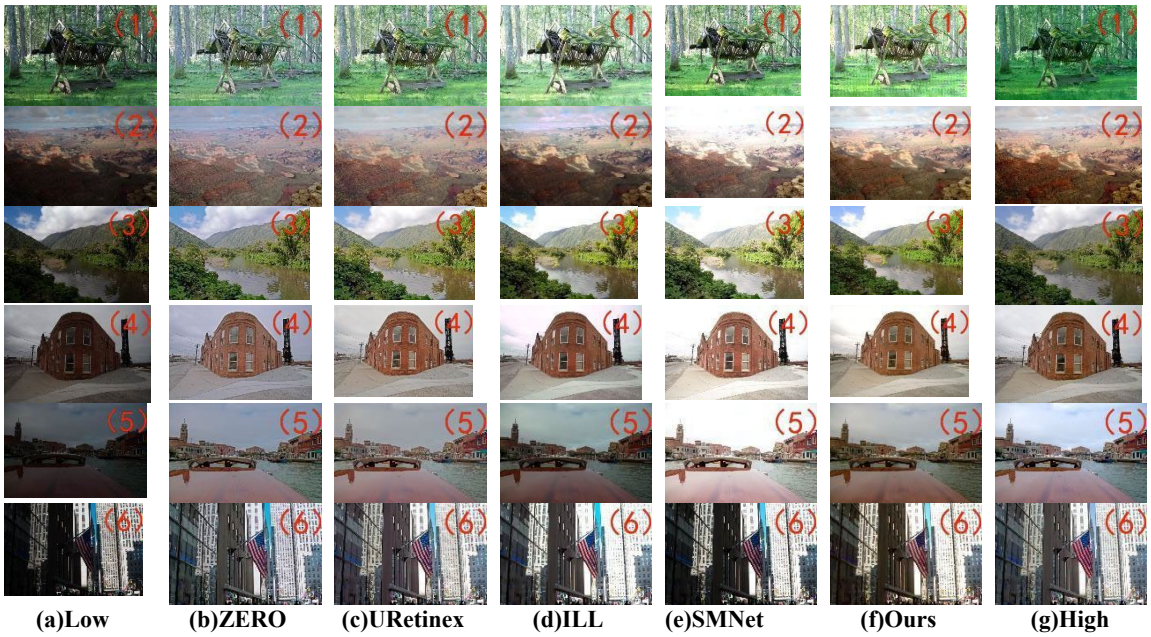



*Figure 8. Restoration results of different algorithms on the GladNet-Dataset*

The test set selects images under outdoor natural lighting conditions to verify the model's generalization ability. The analysis found that the ZERO method results in an overall cool color tone and unclear shadows, with issues of color bias and contrast degradation, as shown in the mountain area of Figure (b)(2), where the color bias is evident. The URetinex algorithm enhances brightness but does not fully recover image details and depth, leading to inadequate detail restoration in the image, as seen in the mountains and forests in Figures (c)(2) and (c)(3), which appear relatively flat and lack a sense of three-dimensionality. The ILL algorithm performs poorly with extremely low-light images, with uneven lighting and some areas underexposed. From the mountains in Figure (d)(2), it can be seen that the brightness recovery of the backlit mountain is insufficient, and the low-light enhancement effect is not ideal. The SMNet algorithm struggles to balance the brightness and contrast of different areas in complex scenes, leading to overexposure or underexposure in certain areas, as shown in Figure (e)(2), where the sky area is overexposed, and Figures (e)(4) and (e)(5) also have overexposure issues. Upon observation and comparison, the images processed by the proposed algorithm have even brightness distribution, reasonable image colors, and well-preserved image details, being closest to the target images overall. This also indicates that the proposed algorithm performs well on images under outdoor natural lighting conditions, further demonstrating the algorithm's generalization ability.

Figure 9 shows the restoration results of different algorithms on the LSRW dataset.

By comparing the images, it was found that the ZERO-DCE algorithm improved over the original low-light images by retaining some details, but the overall brightness enhancement was insufficient, and the details and colors were not rich enough, as seen in the photo wall section of Figure (b)(1). The URetinex algorithm performed well in detail preservation and noise suppression, but the results often had low saturation, as shown by the pink keyboard in Figure (c)(4), with an overall cool color tone and unclear shadow distinction. The ILL algorithm did not perform well with extremely low-light images, and the overall low-light enhancement effect was not significant, as evidenced by the keyboard in Figure (d)(4), which had lost details. The SMNet algorithm, through spatial and multi-scale processing, showed better performance in detail and texture, but introduced some noise and artifacts during image processing, such as the artifacts near the bottle in Figure (e)(6), and the overall image tone was dark. Upon observation and comparison, the images processed by the proposed algorithm demonstrated better results in terms of brightness enhancement, contrast improvement, and detail preservation, being closest to the target images overall, further proving the model's generalization capability.

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207


*Figure 9. Comparison of restoration results of 5 different algorithms on LSRWhuawei dataset*

In addition, peak signal-to-noise ratio (PSNR) and structural similarity index (SSIM) were used to quantify the accuracy of each type for image quality assessment. The higher the PSNR, the better the image quality and the less distortion. The closer the SSIM value is to 1, the better the image quality is and the closer it is to the original image. Universal Image Quality Index (UQI) is an image quality assessment method based on information theory, which takes into account the local mean, variance and correlation of images. The closer the UQI value is to 1, the better the image quality. Learning Perceptual Image Block Similarity (LPIPS) is an image quality evaluation index based on deep learning. The lower the value, the better the image quality and the closer to the original image. Among them, UQI and LPIPS are more focused on the overall quality of the image and human visual perception. Specific results are shown in Table 2, Table 3 and Table 4.

*Table 2. Comparison of objective indicators of different algorithms on the LOL dataset*

| Model | PSNR | SSIM | UQI | LPIPS |
|---|---|---|---|---|
| Guo-ZERO | 14.327 | 0.736 | 0.694 | 0.244 |
| Wu- URetinex | 18.423 | 0.794 | 0.856 | 0.301 |
| Cui-ILL | 21.382 | 0.808 | 0.827 | 0.180 |
| Lin-SMNet | 20.510 | 0.834 | 0.876 | 0.149 |
| **Ours** | **22.133** | **0.873** | **0.907** | **0.133** |

*Table 3. Comparison of objective indicators of different algorithms on the GladNet-Dataset*

| Model | PSNR | SSIM | UQI | LPIPS |
|---|---|---|---|---|
| Guo-ZERO | 18.398 | 0.746 | 0.736 | 0.136 |
| Wu- URetinex | 19.136 | 0.755 | 0.802 | 0.143 |
| Cui-ILL | 18.143 | 0.784 | 0.829 | 0.153 |
| Lin-SMNet | 17.138 | 0.722 | 0.817 | 0.141 |
| **Ours** | **20.427** | **0.807** | **0.846** | **0.124** |

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

*Table 4. Comparison of objective indicators of different algorithms on the LSRW dataset*

| Model | PSNR | SSIM | UQI | LPIPS |
|---|---|---|---|---|
| Guo-ZERO | 16.404 | 0.776 | 0.718 | 0.257 |
| Wu- URetinex | 20.144 | 0.805 | 0.804 | 0.255 |
| Cui-ILL | 19.192 | 0.822 | 0.836 | 0.228 |
| Lin-SMNet | 18.242 | 0.752 | 0.794 | 0.231 |
| **Ours** | **21.242** | **0.826** | **0.868** | **0.225** |

As shown in Table 2, the PSNR value of the LOL test set processed by the proposed algorithm in this paper is slightly lower than the ILL algorithm proposed by Cui et al., and is superior to other mainstream comparison algorithms in the table. This may be due to the extremely low lighting in the low-light images of the LOL dataset, where the performance of the proposed algorithm is not as good as the ILL algorithm. UQI and LPIPS focus more on the overall image quality and human visual perception. The proposed algorithm achieves an SSIM value and UQI value close to 1, while the LPIPS value is also relatively low. This indicates that compared to the other four algorithms, the images processed by the proposed algorithm are more natural and in line with human visual characteristics.

## 5. Conclusions

Compared with the other four methods, the model proposed in this paper achieves excellent results with PSNR and SSIM values reaching 22.133dB and 0.873 on the LOL dataset simultaneously. Additionally, the PSNR on the LSWR low-light image restoration reached 21.242dB. The results demonstrate that the model proposed in this paper not only enhances the brightness and contrast of images but also preserves their natural colors and details, thus achieving better performance in low-light image enhancement tasks. However, the performance of the proposed algorithm on extremely low-light datasets still needs improvement. To address this issue, I will continuously adjust and optimize to achieve the best results.

## References

*[1] Y. Chen. Research on image denoising algorithm based on image filter optimization[D]. Southeast University, 2020.*
*[2] Z. F. Li, L. J. Gao, T. H. Zhang. Simulation of Multi-Feature Detail Recognition of ExposedImage Based on Filtering Algorithm[J]. Computer Simulation, 2023, 40 (12): 242-246+395.*
*[3] Y. C. Hou, H. Song. Remote Sensing Image Denoising Based on WaveletTransform and Improved DnCNN[J]. Computer Measurement & Control, 2022, 216-221.*
*[4] T. Wang, C. Zhao, Z. Z. Chen, et al. Research on denoising method of high frequency ground wave radarocean echo spectrum based on DnCNN network[C]// Chinese Institute of Electronics. Proceedings of the 18th National Annual Conference on Radio Transmission. School of Electronic Information, Wuhan University, 2023: 4.*
*[5]Wu Y ,Xu Q ,Zhang Z , et al. SAR Change Detection Algorithm Combined with FFDNet Spatial Denoising [J]. Journal of Environmental & Earth Sciences, 2023, 5 (2): 1-15.*
*[6] Y. Guo, R. L. Liang, R. M. Wang. Cross-Domain Adaptive Object Detection Based on CNN Image Enhancement in Foggy Conditions[J]. Computer Engineering and Applications, 2023, 59(16): 187-195.*
*[7] X. Z. Yang, W. Wu, T. S. Ren, et al. Complex Environment License Plate Recognition Algorithm Basedon Improved Image EnhancementandCNN[J]. Computer Science, 2024, 51 (S1): 574-580.*
*[8] W. T. Jiang, Y. F. Bu. Image Denoising Network Fusing with CNN and Transformer[J]. Computer Systems & Applications, 2024, 33(07): 39-51.*
*[9] C. Li, C. Guo, and C. C. Loy, "Zero-Reference Deep Curve Estimation for Low-Light Image Enhancement," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2020, pp. 1777-1786.*
*[10] W. Wu, J. Weng, P. Zhang, X. Wang, W. Yang and J. Jiang, "URetinex-Net: Retinex-based Deep Unfolding Network for Low-light Image Enhancement," 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), New Orleans, LA, USA, 2022, pp. 5891-5900.*
*[11] Saputra W V, Suciati N, Fatichah C. Low Light Image Enhancement in License Plate Recognition using URetinex-Net and TRBA[J]. Procedia Computer Science, 2024, 234,404-411.*

ISSN 2706-655X Vol. 7, Issue 2: 46-55, DOI: 10.25236/IJFET.2025.070207

*[12] S. Lin, F. Tang, W. Dong, X. Pan and C. Xu, "SMNet: Synchronous Multi-Scale Low Light Enhancement Network With Local and Global Concern," in IEEE Transactions on Multimedia[J], vol. 25, pp. 9506-9517, 2023.*
*[13] Jongho L ,Man Y R .Dual-Branch Structured De-Striping Convolution Network Using Parametric Noise Model[J].IEEE ACCESS,2020,8, 155519-155528.*
*[14] Y. Chen, X. Dai, M. Liu, D. Chen, L. Yuan and Z. Liu, "Dynamic convolution: Attention over convolution kernels", Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., pp. 11030-11039, 2020.*
*[15] Z. Chen, Z. He, and M. Z. Lu, "DEA-Net: Single Image Dehazing Based on Detail-Enhanced Convolution and Content-Guided Attention," IEEE Transactions on Image Processing, vol. 33, pp. 1002-1015, 2024.*
*[16] Liu, Q., Zhao, J., Cheng, C., Sheng, B., Ma, L.: Pointalcr: adversarial latent gan and contrastive regularization for point cloud completion. Vis. Comput. 38, 3341-3349,2022.*